

IA-00091



Christoph Busch
Fraunhofer Institute of Computer Graphics

Klara Nahrstedt
University of Illinois at Urbana-Champaign

Ioannis Pitas
University of Thessaloniki

Image Security

Over the past decades, research in security has concentrated on the development of algorithms and protocols for encryption, authentication, and integrity of textual data or data with similar characteristics. Despite tremendous advances in security—specifically, the development of asymmetric cryptographic protocols and the inception of strong symmetric ciphers—plenty of security problems still afflict systems. For example, hackers exploiting weaknesses in other systems and the use of inadequate (too short) cipher keys produce frequent news headlines about broken security systems.

Despite the news headlines, such problems have been well explored and even solved in principle, therefore they aren't the primary focus of this special theme issue. Rather, the articles here cover the unsolved problems in image security, which relate fairly closely to computer graphics. The unsolved challenges arise from the increased availability and distribution of multimedia content over Internet services such as the World Wide Web and their implications for intellectual property protection and copyright issues.

A growing number of scientific groups in computer science and cryptography have confronted these challenges. Researchers are currently working on issues such as visual cryptography, mechanisms for the integrity of image material, digital signatures for multimedia data, and data hiding techniques. Data hiding, which has achieved the highest popularity, contemplates the crucial needs for protecting intellectual property rights on multimedia content like images, video, audio, and others. These needs demand robust solutions due to the explosion of publicly available multimedia information and the easiness with which this information can be distributed, copied, and modified. Watermarking technology meets these demands and provides a feasible approach to protect against—and prove-illegal copying and redistribution in the digital world.

This special theme issue presents four articles that discuss watermarking solutions for dedicated media

types such as images, video, and geometric models. They range from an overview of fundamental watermarking concepts to the latest research results.

In "A Watermarking Framework for Copyright Protection of Digital Images," George Voyatzis and Ioannia Pitas review fundamental watermarking concepts and develop a generic model for protecting the copyrights of digital products including a trusted registration authority.

"Digital Watermarking: From Concepts to Real-Time Video Applications" by Christoph Busch, Wolfgang Funk, and Stephen Wolthusen focuses on the specific demands and real-time requirements for video protection. Their solution, which is robust against strong MPEG-2 compression, is very important to the broadcasting and video stock industries.

Boon-Lock Yeo and Minerva Yeung address a new area of watermarking via 3D polygonal models with "Watermarking of 3D Objects for Verification." Their article takes an essential step into this untouched area. They provide mechanisms for embedding watermarks into geometric models as well as detection mechanisms for unauthorized modifications.

"Geometry-Based Watermarking of 3D Models" by Oliver Benedens presents fundamental progress in the same field. His contribution formulates a new definition for robustness of watermarks in correlation to 3D models, proposes a watermarking mechanism, and evaluates its robustness against a polygon simplification.

Finally, we would like to mention that we could not consider all the outstanding submissions received for this theme issue because of page limitations. Contributions were reviewed by up to six security experts from all over the world. We would like to thank all the reviewers, who provided constructive comments and strongly influenced the selection. Their choices allow us to provide a high-quality overview of the state-of-the-art in the field and describe exciting new developments and technologies. We hope you will enjoy reading the following pages. ■

Form SF298 Citation Data

Report Date <i>("DD MON YYYY")</i> 01011999	Report Type N/A	Dates Covered (from... to) <i>("DD MON YYYY")</i>
Title and Subtitle Image Security		Contract or Grant Number
		Program Element Number
Authors		Project Number
		Task Number
		Work Unit Number
Performing Organization Name(s) and Address(es) Fraunhofer Institute for Computer Graphics		Performing Organization Number(s)
Sponsoring/Monitoring Agency Name(s) and Address(es)		Monitoring Agency Acronym
		Monitoring Agency Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited		
Supplementary Notes		
Abstract		
Subject Terms		
Document Classification unclassified		Classification of SF298 unclassified
Classification of Abstract unclassified		Limitation of Abstract unlimited
Number of Pages 10		

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 074-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 1/1/99	3. REPORT TYPE AND DATES COVERED Article	
4. TITLE AND SUBTITLE Image Security			5. FUNDING NUMBERS	
6. AUTHOR(S) Christoph Busch, Klara Nahrstedt, Ioannis Pitas				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) IATAC Information Assurance Technology Analysis Center 3190 Fairview Park Drive Falls Church VA 22042			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Defense Technical Information Center DTIC-IA 8725 John J. Kingman Rd, Suite 944 Ft. Belvoir, VA 22060			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT			12b. DISTRIBUTION CODE A	
13. ABSTRACT (Maximum 200 Words) Over the past decades, research in security has concentrated on the development of algorithms and protocols for encryption, authentication, and integrity of textual data or data with similar characteristics. Despite tremendous advances in security-specifically, the development of asymmetric cryptographic protocols and the inception of strong symmetric ciphers-plenty of security problems still afflict systems. types such as images, video, and geometric models. They range from an overview of fundamental water-marking concepts to the latest research results.				
14. SUBJECT TERMS Image security, security, cryptography			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT None	



Christoph Busch is head of the Department of Security Technology for Graphics and Communication Systems at the Fraunhofer Institute for Computer Graphics, where he is responsible for the acquisition, management, and control of various applied research and development projects. He is also a lecturer on applied wavelet transforms in the educational program of the Computer Graphics Center. He has participated in a series of projects with the Deutsche Telekom AG and Mitsubishi Corp., and is currently a partner in several European projects, including ACTS's Talisman and Octalis projects and Esprit's Aimedia and Filigrane project, all of which deal with copyright protection and conditional access for interactive multimedia services.

Busch studied geodetic sciences at the Technical University of Darmstadt, where he received a PhD in computer graphics in 1997.

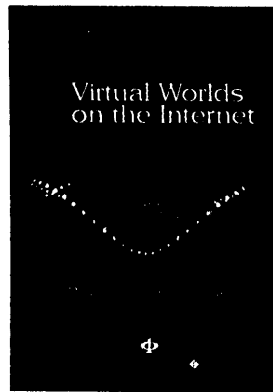


Klara Nahrstedt is an assistant professor in the Computer Science Department at the University of Illinois at Urbana-Champaign. Her research interests are in quality-of-service (QoS) aware resource management for distributed multimedia systems and multimedia security. She received a BA in mathematics from Humboldt University, Berlin in 1984 and an MS in numerical analysis from the same university in 1985. In 1995 she received a PhD from the University of Pennsylvania in the department of Computer and Information Science.



Ioannis Pitas is a professor at the Department of Informatics at the University of Thessaloniki. His current research interests are in the areas of digital image processing, multidimensional signal processing, and computer vision. He received the diploma of electrical engineering in 1980 and a PhD in electrical engineering in 1985 at the University of Thessaloniki, Greece. He co-authored the book *Nonlinear Digital Filters: Principles and Applications* (Kluwer, 1990) and authored *Digital Image Processing Algorithms* (Prentice Hall, 1993). He is currently an associate editor of IEEE Transactions on Neural Networks.

Readers may contact Busch at Fraunhofer Institut Graphische Datenverarbeitung, Rundeturmstrasse 6, D-64283 Darmstadt, Germany, e-mail busch@igd.fhg.de. Contact Nahrstedt by e-mail at klara@cs.uiuc.edu and Pitas at pitas@zeus.csd.auth.gr.



Virtual Worlds on the Internet

John Vince and Rae Earnshaw

Virtual Worlds on the Internet examines how the latest developments in virtual environments, animation, communication networks, and the Internet are configured to create revolutionary tools and systems. Vince and Earnshaw selected twenty papers they believe will influence computer systems of the

twenty-first century. These papers illustrate topics such as a toolkit for the development of virtual environment applications, different uses of VRML in information system interfaces, an examination of research in virtual reality environment interfaces, and five approaches to supporting changes in virtual environments.

360 pages. 7" x 10" Softcover. January 1999. ISBN 0-8186-8700-2
Catalog # BP08700 — \$44.00 Members / 155.00 List

Online Catalog

<http://computer.org>

+1 800.CS.BOOKS



Information Visualisation IV99

International Conference
on
Information Visualisation
· LONDON · ENGLAND

14-16 July 1999

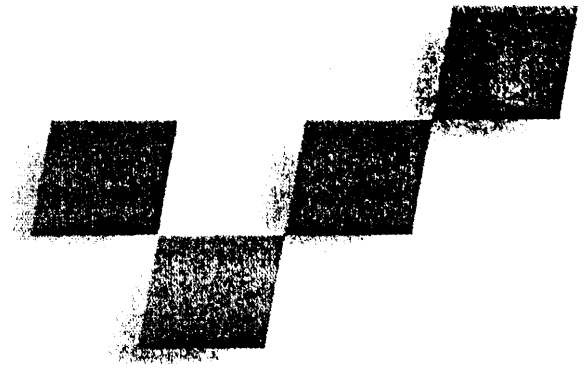
Call for Papers, Videos and
Participation

For more information see the
conference web site:

<http://www.it-link.demon.co.uk/IV99/>

Or contact Ebad Banissi at
E.banisse@sbu.ac.uk

Protecting Digital-Image Copyrights: A Framework



George Voyatzis and Ioannis Pitas
University of Thessaloniki

In the past decade, the scientific community has engaged in an intensive discussion of security issues for digital data. Piracy of such data has obstructed the rapid evolution of digital networks, digital libraries, and World Wide Web services. The convenient broadcasting or exposition of digital products on the global network leads easily to illegal copying and retransmission. The same happens with data transmitted through insecure channels.

Today, digital data security covers such topics as access control, authentication, and copyright protection for still images, audio, video, and multimedia products. The networking environment of the future will require tools that provide

- secure and fast digital data encryption and decryption,
- content verification (authentication) of the received data by the recipient, and
- robust and trustworthy marks indicating copyright and legal ownership.

Digital watermarks offer a way to counter copyright piracy on the global network. We summarize the fundamental concepts of watermarking and describe a general framework for a copyright protection system.

For the first of these items, private- and public-key cryptography are widely used and have well-established algorithms. (These include RSA, the Rivest-Shamir-Adleman algorithm, and DES, the Data Encryption Standard?) For authentication, one proposal is digital signatures, which are also based on cryptographic algorithms? In this article, we address the third item. We describe how *digital watermarking* can contribute significantly to a realistic framework for protecting intellectual property rights in digital media.

What is a digital watermark?

The concept of the digital watermark is a new topic^{4,5} associated with the data-hiding technique known as steganography? Steganography, or “covered writing,”

has a long history that includes various methods of secret communication. In contrast to cryptography, **steganography** does not immediately arouse suspicion of something **secret** or valuable. Instead, it hides an important message within an unimportant one. **Watermarking**—also called tamper-proofing or content **verification**⁷—hides a secret and personal message to protect a product’s copyright or to demonstrate its data integrity.

An important difference between steganography in general and watermarking relates to the attacker’s goal. A pirate tries to reveal the information a steganography message carries. However, a pirate tries either to **remove** a watermark to violate a copyright or to cast the same watermark, after altering the data, to forge proof of authenticity.

Generally, the watermarking of still images, video, and audio demonstrate certain common fundamental concepts. Without significant loss of generality, we will focus on watermarking still images with a sufficient number of grayscale levels or colors (**8-bit** grayscale or **24-bit** color images). Using such images as examples, we discuss the efficiency of watermarks for copyright protection.

Digital image protection via watermarking

To understand the role of watermarks in an intellectual property (IP) protection framework, consider the elementary digital media delivery system presented in Figure 1. The user gets a digital product from a provider—the copyright owner or an authorized distributor. Access and transmission take place in a global network environment. Users who do not retransmit the products do not violate the copyright property. Pirates harm the copyright owner by reproducing and retransmitting digital products illegally. That is, pirates function as unauthorized providers.

A protection scheme should give providers a reliable method for efficiently searching the network for copies that originally belonged to them. Such a scheme should also supply strong indications for asserting legal ownership.

A digital watermark (called also a copyright label or

invisible stamp) is associated with a unique private identification number called a *watermark key*. Under a watermarking-based protection scheme, each provider possesses a unique secret **key** K_{pr} (or a finite set of keys). The scheme itself consists of the following **actions**:^{1,7}

- The copyright owner alters the digital data of the original image I_0 to produce the watermarked image I_w by using the private key and a public or private algorithm $\hat{E}: I_w = \hat{E}(I_0, K_{pr})$.
- The copyright **owner** can detect the watermark by using an algorithm \hat{D} :

$$\hat{D}(I, K) = \begin{cases} 1 & \text{if } I \sim I_w \text{ and } K = K_{pr} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

The relation $A \sim B$ denotes perceptual similarity between images A and B . $I_0 \sim I_w$ because watermarking should produce almost invisible alterations.

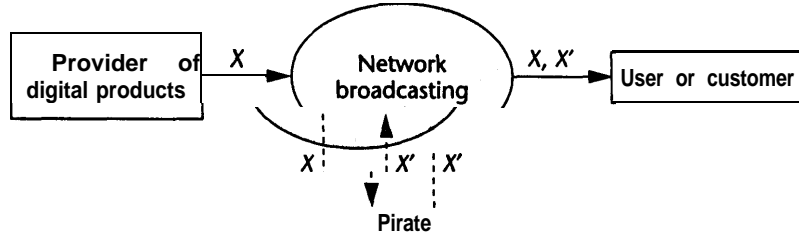
To complete the watermarking scheme, we include the following actions, which are optional on the part of the provider:

- The provider possesses a personal archive L of original digital images and a matching procedure \hat{m} such that

$$\hat{m}(I, L) = \begin{cases} 1 & \text{if } \tilde{I} \in L \text{ and } \tilde{I} \sim I \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

- The provider registers watermark **key** K_{pr} to a trusted authority and ensures its uniqueness.
- The provider registers watermarked images $I \in L$ with a trusted authority to establish (time-stamp) ownership of I .

The aim of this copyright protection **system** is to enable providers to check possible violation of **their** copyright property by monitoring the traffic and exposition of digital images in the distribution network. Positive detection of the provider's watermark ($\hat{D}(I, K_{pr}) = 1$) is an indication of ownership that the provider can use as substantiated evidence in a court of law. The registration of the originals to a trusted authority can give full proof of ownership in case of detection ambiguities—for example, if there has been an attack on the watermark.



1 Basic system for digital image distribution.

Basic watermarking procedures

The literature contains various definitions of digital **watermarks**.⁸⁻¹⁰ For digital images of size $N \times M$, it is convenient and quite general to consider 2D watermark signals that consist of binary or, more generally, ternary elements:

$$w = \{w(k) : w(k) \in \{-1, 0, 1\}, k \in \hat{W}\}$$

or white noise in $[-1, 1]$. \hat{W} denotes a 2D grid of size $N \times M$. Vector k indicates the grid positions of the watermark pixels.

The three fundamental stages of watermarking are generation, embedding, and detection. Figure 2 shows a schematic presentation of the overall algorithm.

Watermark generation

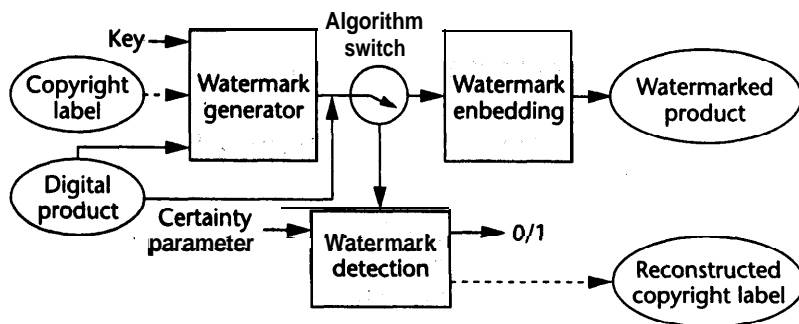
Let W be the set of possible watermark signals—that is, all the signals having elements in $\{-1, 0, 1\}$. We consider the finite and countable space K of keys. If X denotes the set of digital images, we define a watermark generation procedure by the noninvertible function

$$G: X \times K \rightarrow W, W = G(I, K)$$

where $K \in K$ is the watermark key, and $I \in X$ is the image in which the watermark will be embedded. Without loss of generality, we decompose G as follows:

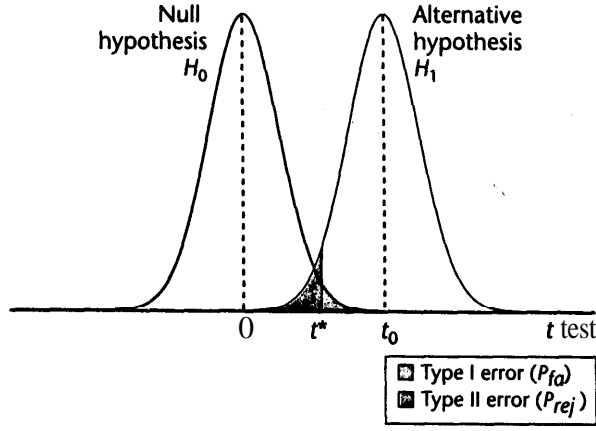
$$G = T \circ R, R: K \rightarrow W, T: W \times X \times K \rightarrow W$$

R should be noninvertible” and may be based on a pseudorandom-number generator or, generally, on chaotic systems. When we insert **key** K in R , we create a fixed watermark $W_0 = R(K)$. T modifies W_0 to obtain the final watermark W that depends on the particular image. The watermark modification function T should take into account only robust characteristics of the image data. The original image I_0 , the watermarked one I_w , and a modified copy of I_w (denoted I'_w) must result



2 Overall watermarking algorithm, including embedding and detection, both of which require watermark generation first. The copyright label is a binary stream that carries public information about ownership.

3 The distributions of the detector output for the null hypothesis (watermark does not exist) and the alternative hypothesis (watermark exists) for statistical detection, based on the t test. Certainty ($c = 1 - P_{fa}$) depends on the selection of the t threshold t^* .



procedure directly to the intensity or luminance domain, we can apply it to the discrete cosine transform (DCT) domain^{8,9} or the wavelet domain.¹²

Watermark detection

Watermark detection is the most crucial part of a watermarking framework. By adopting a hypothesis-testing framework, in practice we obtain detection errors and deviate from the ideal detection (Equation 1). Such errors fall into two categories:

- **Type I errors.** We detect the watermark, but it does not actually exist in the data. Such false positives contribute to the probability of false alarm (P_{fa}).
- **Type II errors.** We do not detect the watermark in the data, but it does exist. Thus, we get false negatives that contribute to the probability of false rejection (P_{rej}).

Taking these error probabilities into account, we define a detection procedure as follows:

$$D: \mathbf{X} \times \mathbf{K} \rightarrow (0, 1) \subset \mathbb{R}$$

$$D(I, W) = D(I, G(I, K_{pr})) = c$$

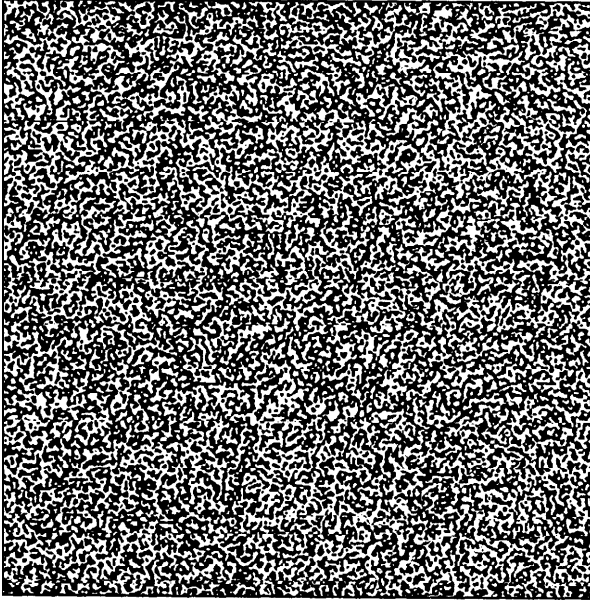
The output c denotes the certainty of a positive detection. The final decision is based on

$$c \geq c_{thres} \Rightarrow \text{watermark exists}$$

Generally, $c = f(t)$, where t is a test statistic. The parameter c_{thres} is the certainty level for detection, chosen by the image provider who applies detection. It is associated with a test value t that controls (balances) the two error types. Figure 3 illustrates this situation. When c is close to unity, false negatives become insignificant ($P_{fa} \rightarrow 0$). However, at the same time, P_{rej} increases and, consequently, the detector's overall performance declines. D approaches the ideal detector \hat{D} (see Equation 1) when both P_{fa} and P_{rej} tend to zero.

This detection scheme performs a binary decision with a given certainty level and answers the fundamental question, "Does watermark $W(I, K_{pr})$ exist in product I ?" In the framework we propose, only the legal owner, who possesses key K_{pr} , can perform detection. Therefore, in principle, it is redundant to hide multi-bit information identifying the owner in the watermarked product. However, bitstream information hiding and retrieval (possibly error-corrected¹³) can increase certainty in the detection process or give illustrative indications of ownership (such as a company logo) or store product manipulation operations. It would also be useful to embed ownership information for other reasons—for example, if a trusted authority watermarks the products of various owners with the same key.

4 A chaotic binary watermark for embedding in a digital image.



adequately in the same watermark when perceptual similarity is satisfied:

$$w = T(W_0, I_0) \cong T(W_0, I_w) \cong T(W_0, I'_w) \\ \text{for any } I_0 \sim I_w \sim I'_w$$

Watermark embedding

We define the embedding procedure as a superposition of the digital watermark signal $W = \{w(k)\}$ onto the original image $I_0 = \{x_0(k)\}$ of the same dimension. We denote the embedding procedure by E , which we define as

$$E: \mathbf{X} \times \mathbf{W} \times \mathbb{R}^2 \rightarrow \mathbf{X}, x_w(k) = x_0(k) \oplus H(k) \otimes w(k) \quad (3)$$

$H(k)$ is a 2D watermark embedding mask that controls watermark visibility and robustness. \oplus is a superposition operator including appropriate truncations and quantization, and \otimes is a generalized multiplication operator. The expression $I_w = \{x_w(k)\}$ denotes the watermarked image. The recovery of the original from the watermarked image is desirable but not necessary, because the copyright owner can store the original in a private archive. Instead of applying this embedding

A watermarking example

Several already proposed watermarking schemes would work with the general framework we describe; we have presented such a scheme previously.¹⁴ For the purposes of illustration, we summarize a simple version of the method and apply it to an image.

Step A. The bivalued watermark $W_0 = \{w(i, j)\}$, presented in Figure 4 as a black-and-white image, is created by watermark generation procedure G using a chaotic system. The set of integer numbers is the watermarking key set. Each integer corresponds to a set of parameters for the chaotic system through a well-defined function. Therefore, each key represents a unique initial condition of the chaotic sequence and can create a sufficiently different watermark for each different key.

Using chaotic systems, we can efficiently control the watermark's low-pass characteristics to achieve sufficient robustness under lossy compression or low-pass filtering.

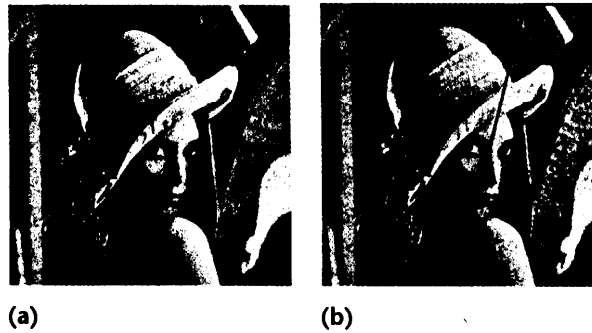
Step B. Let $I_0 = \{x(i, j)\}$ denote the 8-bit luminance space of the original image (Figure 5a) to be watermarked. As we discussed earlier, we apply operator T to W_0 to obtain an image-dependent watermark W . T must preserve W 's main features—in our example, its low-pass character. This is feasible if T alternates (or doesn't alternate) the watermark values by taking into account robust characteristics of the image regions indicated, for example, by vector $t = p + r(K)$. Here, p denotes the particular watermark pixel examined, and r is a key-dependent vector.

Step C. We form the watermarked image I_w , which is the same size as the original, by applying equation 3 and using a constant-embedding mask ($h_{ij} = h = \text{const}$). The symbol \oplus denotes the usual addition operator and the necessary truncations.

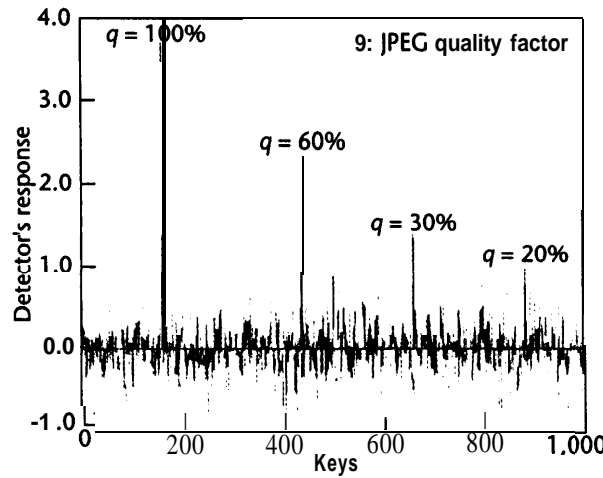
$$xx_w(i, j) = \begin{cases} 0 & \text{if } x(i, j) + hw(i, j) < 0 \\ 255 & \text{if } x(i, j) + hw(i, j) < 255 \\ x(i, j) + hw(i, j) & \text{otherwise} \end{cases}$$

The constant h should be small to guarantee the invisibility of the watermark in I_w . For $h = 2$ the result is the watermarked image presented in the left part of Figure 5b. When we use $h = 8$, as in the right part of Figure 5b, the watermark embedding causes clearly visible degradations.

Step D. We can apply detection to any image I of the same size. We generate watermark W using the specific key and the image I . The image pixels are divided into two



5 Embedding a digital watermark. (a) The original image, "Lenna." (b) Invisible watermarking of "Lenna" using $h = 2$ on the left section and $h = 8$ on the right section. The watermarking on the right section degrades the image.



6 Watermark detector output for various incorrect keys (red peaks) and correct keys (blue peaks). With the correct keys, the system easily distinguishes the watermark in the original watermarked image and high-quality compressed versions of it.

subsets $A = \{(i, j) | w_{ij} = 1\}$ and $B = \{(i, j) | w_{ij} = -1\}$, which have mean luminance values \bar{A} and \bar{B} . We consider the following detection response:

$$R = D(I, W) = \bar{A} - \bar{B}$$

We assume that x and w are statistically independent random variables. Thus, R should follow the normal distribution $N(\bar{R}, \sigma)$, where $\bar{R} = 2h$ when W exists in I and $\bar{R} = 0$ otherwise. When a watermarked image I_w is modified somehow (for example, by lossy compression or filtering), we expect $R < 2h$. We assume the watermark exists when $R > R_{thres}$. The constant R_{thres} is a threshold value associated with a certainty level C_{thres} for a positive detection. After numerical experiments, we found that for $R_{thres} > 0.1$, the minimum detection certainty is 99.9 percent.

Figure 6 shows the detection response R obtained from I_w by using 1,000 incorrect watermark keys. The blue peaks of R come from I_w or versions of it that we JPEG-compressed for various JPEG (Joint Photographic Experts Group) quality factors. We used the correct key for each of these peaks. With a decreased R_{thres} , the system can detect the watermark even with high compression ratios (or, generally, major modifications). However, this would cause false alarms as well.

Basic watermarking demands

The basic watermarking methods we have presented efficiently protect the copyright property if the following features characterize the elementary procedures G, E, and D.

Perceptual quality preservation

Embedding a watermark in a digital image, video, or audio signal should not reduce its perceived quality. An invisible watermark satisfies this requirement. Therefore, we should take human visual perception into account when casting watermarks. In the embedding scheme described by Equation 3, the proper choice of the embedding mask $H(\mathbf{k})$ guarantees invisibility. That is, $|H(\mathbf{k})| < \text{JND}$, where JND denotes just-noticeable distortion of the image. Some watermarking schemes have used human visual models that provide strong but invisible image alterations.^{15,16}

Generally, however, watermark invisibility implies potential watermark removal by lossy compression, and, possibly, by other processing operations. Small perceptible alterations that cannot be considered distortions of the original, or that do not reveal artifacts under postprocessing operations (if necessary), offer a possible alternative. We stress that, in most cases, only the watermarked image is available to the user.

Trustworthy detection

We can evaluate the performance of a watermark detection procedure D by using the total error probability $P_{\text{err}} = P_{\text{fa}} + P_{\text{rej}}$. The contribution of the errors P_{fa} and P_{rej} to form the total detection error P_{err} is controlled by certainty level c_{thres} . We distinguish the following cases:

- Detection **with low** certainty. False alarms are frequent. However, the probability of missing the detection of a watermarked image is very small. Providers can ascertain the reliability of a positive detection by searching their archive for the image and applying the matching process \hat{m} .
- Defection **with high** certainty In this case, $P_{\text{fa}} \rightarrow 0$, and the detector provides very reliable positive detection. Watermark detection with extremely small false-alarm probability may stand by itself as sufficiently reliable evidence for proving legal ownership in a court of law. However, the rejection probability may become quite significant.

The capability of G to produce an enormous watermark set W consisting of obviously distinguishable elements in the detection procedure is a necessary condition for achieving very low false-alarm probabilities. Therefore, rich complexity should characterize watermark signals (for example, see Figure 4).

Computational efficiency

Watermark casting and detection are two separate algorithms. They should be easy to use, applicable to any digital image, and trustworthy. Although the provider embeds watermarks in each product only once, detection should apply to a large set of publicly exposed images. Therefore, very fast watermark detection is desirable. The private key and the watermarked image should suffice to perform watermark detection. Watermarks can be combined with an automated search procedure that searches for illegal copies in any public network or broadcasting environment (web-crawling or monitoring).

A necessary condition for development of such a system

is to avoid the use of the original product (or any information associated with it) in the detection algorithm. Otherwise, the network monitor would have to apply matching procedure \hat{m} (Equation 2) between all network images examined and all images in the private archive. This would result in unacceptable computational complexity and make watermarking entirely useless. In our approach, watermarking is essentially a copyright signaling mechanism.

Robustness to digital processing

Watermark robustness under image modification is an essential issue for copyright protection.^{8,9} Any provider or user can modify an original digital image to improve quality, compress data, edit digitally, and so on. Protecting copyright while maintaining sufficient quality under these conditions is desirable. Preservation of image quality implies preservation of perceptual similarity. Usually, we demand watermarks to be robust under the following image-processing operations:

- **Lossy compression.** It is the most widely used procedure for storing and transmitting digital still images, video, and audio. The widely used JPEG and MPEG (Moving Pictures Experts Group) algorithms provide a high compression ratio and the desired quality. However, compression algorithms tend to remove invisible information that can be related to the watermark. Therefore, watermarks should combine invisibility and robustness simultaneously.
- **Filtering and enhancement.** Users apply filtering to remove noise or to improve the perceptual quality. This process can remove watermarks as well. In addition, attackers may develop filters specifically designed for watermark removal. Denoising filters are usually low-pass; thus, watermarks should possess low-pass characteristics.
- **Geometric modifications** In the case of digital images such modifications include scaling, rotation, cropping, reflection, line and column extraction or insertion, and combinations of these. Salient image features or the invariant properties of the Fourier domain have been used for casting robust watermarks? The watermark should cover the entire image so that it will be robust to cropping. Watermark detection in geometrically modified products without resorting to the original product is a difficult task. Although there has been considerable progress in this area recently, the problem is still unsolved in its general form.”
- **Different image presentations.** A user or pirate can print a watermarked digital image and then create a new digital copy by rescanning. If this process does not reduce the image quality significantly, the watermark should still be detectable in the rescanned image.
- **Color correction.** Usually, watermarks are embedded in the luminance component so that they are resistant to compression. Therefore, they are resistant to color corrections.

Intentional attacks on watermarks

The watermarking scheme we have described may be attacked directly or indirectly to undermine its capaci-

ty to indicate legal ownership.^{18,19} The following sections discuss possible attacks and methods of defense against them. Remember, our scheme does not allow the use of originals in the detection procedure.

Extraction of counterfeit watermarks

Pirates know the principle of the public **watermark**-ing detection algorithm. Thus, they may form a signal **W** for a particular image **I**, which forces the detector to indicate that the image is watermarked? In other words, **W** is a watermark signal never embedded in **I**, but the pirate uses it as his or her own watermark. However, **W** should be associated with a **key** $K \in K$ through the watermark generation procedure **G**. Because this procedure is generally not invertible, such an attack is extremely unlikely to succeed. Procedure **R** usually inherits **G**'s **non**-invertibility. Qiao and Nahrstedt describe a **noninvertible** scheme based on the DES **cryptosystem**.¹¹

Detection of false positives

A pirate, after a trial-and-error procedure, might find a key **K** that provides a positive detector response and then claim that **K** indicates his or her ownership. However, the pirate can find such a key only after approximately $1/P_{fa}$ applications of **D**, where P_{fa} is the minimum acceptable false-alarm probability. Watermark detection with relatively significant P_{fa} becomes quite reliable when providers register their keys with a trusted authority. In such a case, the probability that the counterfeit key **K** coincides with a provider's registered key is extremely small.

Statistical watermark extraction

A great number of digital images, all watermarked with the same key, must not reveal the watermark when the pirate applies statistical methods such as averaging. Therefore, watermarks should not be statistically recoverable. An efficient protection against such an attack is to use image-dependent **watermarks**.²⁰ We have anticipated this requirement by introducing procedure **T** during watermark generation.

Multiple watermarking

Multiple watermarking can be useful when product resellers want to embed their own watermark in a **product**.^{9,10} However, an attacker may use this property to embed his or her own watermark. Both the original and the pirate's watermarks can be detected using the corresponding unique key. Generally speaking, we can detect the watermark embedded at a later stage with greater certainty. However, in general, we cannot definitely conclude which watermark was embedded first. In this case, the image's original owner is the only one who can produce a copy containing only his own watermark. Subsequently, the original owner can prove legal ownership in cases of conflict. This eventuality explains the need for a private archive and watermarked product registration, as described earlier.

Watermarking with arbitrary keys

Malevolent users or pirates may apply watermarks to any accessible image by using arbitrary keys—a process

known as “watermark bombing.” If these products are publicly exposed, great confusion may arise during the automated watermark-searching procedure. The negative consequences of such an indirect attack are restricted by the fact that we demand an enormous set of watermark keys. An efficient method of preventing such an attack is to consider distribution restrictions on provider keys. Each watermarking software package should supply just one or a few keys to each provider for watermark casting.

Private key loss or theft

The loss of the private key can enable a pirate to remove the watermarks from all the images that belong to that particular owner. This would make the system dangerously unstable. A solution to this serious problem could be provided by a combination of secure **time-stamping**² and time-dependent watermarks. These actions would restrict such an attack to a small number of products. This solution has the drawback of increased detection complexity.

Miscellaneous modifications

A pirate can apply a sequence of various uncommon image-processing operations to confuse the monitoring software or to desynchronize the detector. (For example, the pirate might use the “mosaic attack,” which is essentially a cropping **attack**.¹³) Unfortunately, once the data are out in the distribution network, there is always the risk of watermark removal by new techniques. In this case, the sole protection is the registration procedure.

Product registration

The use of Web servers or trusted third parties in a copyright protection system is one proposal for solving the problems that exist in a pure watermarking technique.¹¹ Product registration to a trusted authority is a well-established way of protecting intellectual property rights of various products—for example, books, software packages, and so on. Registration information can offer indisputable proof of original ownership and legal rights. A protection system based on product registration requires the following actions before product disposition:

1. The provider registers with a trusted authority, which provides a watermarking key (or set of keys) or, generally, a private watermark-embedding software package.
2. The provider uses the registered key or the private software to cast watermarks.
3. The provider includes the original product in an archive and registers the watermarked product with a trusted authority.
4. The provider conducts an automated watermark search on the distribution network. A provider who uses low-certainty detection reinforces the **reliability** of a positive detection result by searching the archive.
5. Producing the registered copy in a court of law constitutes proof of copyright ownership.

As this scenario shows, the inclusion of registration

chemes in the protection system means that the watermark contribution is restricted to the monitoring and **discovery** of illegally distributed products. The watermarking detector should provide a reliable detection **certainty** level, recommending that the provider **pro-**ceed with or decline further investigation by searching he archive. Afterwards, the registration authority must **provide** the final and reliable proof of legal ownership.

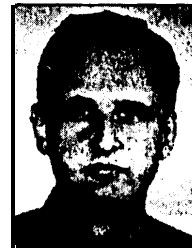
Conclusions

Watermarks efficiently protect copyright when basic demands can be satisfied. However, the demand for **watermarks** to remain robust under digital **image-processing** operations has not yet been fully satisfied. **Robustness** to geometrical distortions is an essential remaining problem for schemes that do not use the original images in the detection procedure. Furthermore, new water-marking schemes must take into account **optimizations** of current processing algorithms and the development of new algorithms. Better future compression techniques, for example, must not destroy watermarks embedded at any earlier time.

A watermarking scheme should also be resistant to intentional attacks. Pirates will try to violate the protection system either by creating forged proofs of ownership or by undermining the capacity of the watermarking scheme to indicate legal ownership. We believe that a general scheme for effectively protecting digital products can be based on a combination of watermarking, the use of private archives, and product registration with trusted registration authorities. ■

References

1. B.M. Macq and J.J. Quisquater, "Cryptology for Digital TV Broadcasting," *Proc. IEEE*, Vol. 83, June 1995, pp. 944-957.
2. D.R. Stinson, *Cryptography Theory and Practice*, CRC Press, New York, 1995.
3. T. ElGamal, "A Public Key Cryptosystem and Signature Scheme Base on Discrete Logarithms," *IEEE Trans. Information Theory*, Vol. 31, No. 4, 1985, pp. 469-472.
4. Special issue on watermarking, *Signal Processing*, Vol. 66, No. 3, May 1998.
5. Special issue, *IEEE J. Selected Areas in Comm.*, Vol. 16, No. 4, 1998.
6. N.F. Johnson and S. Jacodia, "Exploring Steganography: Seeing the Unseen," *Computer*, Vol. 31, No. 2, Feb. 1998, pp. 26-34.
7. W. Bender, D. Gruhl, and N. Morimoto, "Techniques for Data Hiding," *Proc. SPIE*, Vol. 2420, SPIE, Bellingham, Wash., 1995; p. 40.
8. E. Koch and J. Zhao, "Towards Robust and Hidden Image Copyright Labeling," *Proc. IEEE Workshop on Nonlinear Signal and Image Processing*, IEEE, 1995, pp. 452-455.
9. I.J. Cox et al., "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Trans. Image Processing*, Vol. 6, No. 12, 1997, pp. 1673-1687.
10. I. Pitas and T.H. Kaskalis, "Applying Signatures on Digital Images," *Proc. IEEE Workshop on Nonlinear Signal and Image Processing*, IEEE, 1995, pp. 460-463.
11. L. Qiao and K. Nahrstedt, "Watermarking Schemes and Protocols for Protecting Rightful Ownership and Customer's Rights," to be published in *J. Image Comm. and Image Representation*, 1998.
12. D. Kundur and D. Hatzinakos, "A Robust Digital Image Watermarking Method Using Wavelet-Based Fusion," *Proc. ICIP 97*, Vol. I, 1997, pp. 544-547.
13. J.J.K. Ruanaidh and T. Pun, "Rotation, Scale, and Translation Invariant Spread Spectrum Digital Image Watermarking," *Signal Processing*, Vol. 66, No. 3, 1998, pp. 303-317.
14. G. Voyatzis and I. Pitas, "Chaotic Watermarks for Embedding in the Spatial Digital Image Domain," *Proc. ICIP 98*, IEEE, 1998, pp. 432-438.
15. J.F. Delaigle, C. De Vleeschouwer, and B. Macq, "Watermarking Algorithm Based on a Human Visual Model," *Signal Processing*, Vol. 66, No. 3, 1998, pp. 337-355.
16. C.I. Podilchuk and W. Zeng, "Image Adaptive Watermarking Using Visual Models," *IEEE J. Selected Areas in Comm.*, Vol. 16, No. 4, 1998.
17. F. Petitcolas, R.J. Anderson, and M.G. Kuhn, "Attacks on Copyright Marking Systems," *Proc. Second Workshop on Information Hiding*, 1998.
18. S. Craver et al., "Resolving Rightful Ownerships with Invisible Watermarking Techniques: Limitations, Attacks, and Implications," *IEEE J. Selected Areas in Comm.*, Vol. 16, No. 4, 1998.
19. I.J. Cox and J.P. Linnartz, "Some General Methods for Tampering with Watermarks," *IEEE J. Selected Areas in Comm.*, Vol. 16, No. 4, 1998, pp. 587-593.
20. N. Nikolaidis and I. Pitas, "Robust Image Watermarking in the Spatial Domain," *Signal Processing*, Vol. 66, No. 3, 1998, pp. 385-403.



George Voyatzis is a researcher in the Department of Informatics at the University of Thessaloniki, Greece. His research concerns applications of chaotic systems in celestial mechanics and in digital processing. He earned a diploma and PhD in physics at the University of Thessaloniki in 1987 and 1993, respectively.



Ioannis Pitas is a professor at the Department of Informatics at the University of Thessaloniki. His current research interests are in the areas of digital image processing, multidimensional signal processing, and computer vision. He received the diploma of electrical engineering in 1980 and a PhD in electrical engineering in 1985 at the University of Thessaloniki, Greece. He co-authored the book *Nonlinear Digital Filters: Principles and Applications* (Kluwer, 1990) and authored *Digital Image Processing Algorithms* (Prentice Hall, 1993). He is currently an associate editor of IEEE Transactions on Neural Networks.

Readers may contact Pitas at the Department of Informatics, Aristotle University of Thessaloniki, Box 451, Thessaloniki 54006, Greece, e-mail pitas@zeus.csd.auth.gr.